ote technical note techn

Proposed Selection Criteria for Aviation Safety Analytical Methods and Tools

Jacques Press, ACT-560

June 1999

DOT/FAA/CT-TN99/16

Document is available to the public through the National Technical Information Service, Springfield, Virginia 22161

U.S. Department of Transportation
Federal Aviation Administration

William J. Hughes Technical Center
Atlantic City International Airport, NJ 08405

NOTICE

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
DOT/FAA/CT-TN99/16		

4. Title and Subtitle	5. Report Date
Proposed Selection Criteria for Aviation Safety Analytical Methods and Tools	June 1999
	6. Performing Organization Code ACT-560
7. Author(s) Jacques Press, ACT-560	8. Performing Organization Report No. DOT/FAA/CT-TN99/16
9. Performing Organization Name and Address Federal Aviation Administration William J. Hughes Technical Center Atlantic City International Airport, NJ 08405	10. Work Unit No. (TRAIS)
	11. Contract or Grant No.
12. Sponsoring Agency Name and Address	13. Type of Report and Period Covered Technical Note
	14. Sponsoring Agency Code

15. Supplementary Notes

16. Abstract

This report provides a set of criteria useful in choosing and using analytical tools and methods (artifacts) directed at aviation safety analysis. The approach consists of adopting a rational framework in three selection stages: artifact classification, value, and quality. Furthermore, we have supplied a scoring and weighting method as an example. Over the years, experts have devised numerous analytical artifacts to articulate safety data into a comprehensive body of knowledge. Given the seriousness of aviation safety, we believe all such artifacts need to be evaluated wisely and cautiously before any claims are made. The criteria prescribed in this report provide one way to classify and evaluate them consistently, as a prelude to their use.

17. Key Words	18. Distribution Statement
Selection Criteria Analytical Tools and Methods Air Safety	This document is available to the public through the National Technical Information Service, Springfield, Virginia, 22161

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	20	

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

Table of Contents

List of Illustrations

Executive Summary

Over the years, the aviation community has devised many analytical ways (artifacts) to encapsulate meaningfully safety data into a comprehensive body of knowledge. Given the seriousness of aviation safety as a societal issue, we believe all such artifacts should be evaluated wisely and cautiously before claims are made about the results obtained through these tools. Moreover, common sense tells us we should assess objectively the value and quality of these artifacts, given that the safety concept itself remains quite complex to characterize analytically.

Accordingly, this technical note provides a generic set of criteria useful in choosing and using all sorts of artifacts directed at aviation safety analysis. The criteria approach consists of adopting a rational framework in three selection stages: classification, value, and quality. The framework is meant to instill discipline and formality in the selection process. The technical note includes a scoring and weighting method supplied as an example. Using this example, a selection team could then identify and retain valuable artifacts into an informational depository, periodically updating and making its contents available to the aviation community.

1. Introduction

1.1 Background

Safety information plays a valuable role in aviation. Motivated by such a premise, experts are devising numerous analytical methods and tools meant to articulate this information further into an understandable and useful body of knowledge. They know society considers aviation safety an important issue. They also know the aviation community feels compelled to believe that most safety analyses, whether extensive or minimal, bring some sort of added value to the world. Consequent to this belief, analytical artifacts (e.g., methods and tools) continue to proliferate in many safety domains. Despite the trend, common sense tells us we should assess objectively the value and quality of these artifacts, given that the safety concept itself remains quite complex to characterize analytically. Appendix A provides a representative list of such artifacts. The community needs to evaluate rationally all methods and tools so that practitioners can choose and use them wisely and cautiously.

In response to this need, it would be beneficial to compile a list of methods and tools. This proposed list could then be shared with the community through a dissemination initiative. In the long run, the list should to be more than a simple catalogue. It should become a living depository that experts will hopefully enrich, annotating it with attributes and descriptors that qualify each analytical artifact for its potential contribution to aviation safety understanding.

1.2 Purpose

As prelude to the compilation, this document proposes a set of selection criteria to apply to all artifacts. Using the process described in this document, a selection team could then identify and retain the valuable artifacts into an informational depository. We foresee an initial list, with equally initial attribute ratings derived from the selection criteria process. We also foresee someone updating the list periodically, as users' feedback becomes available.

2. Methodology

2.1 Selection Criteria

We propose several criteria designed purposefully in three sequential evaluative stages: classification, value, and quality. Each stage would come equipped with its own scoring process to assess individual analytical artifacts. The stages are summarized as follows:

- Classification (Stage 1) rates each artifact by its (a) aviation safety relevancy, (b) analytical nature, and (c) maturity state. Stage 1 uses a two-way classification based on the framework prescribed in the next subsection. We believe defining artifacts this global way helps us decide at the onset whether to select them or not. Because Stage 1 filters them in such a universal context, we propose that it carries at least half the total score (e.g., 50 points out of the total maximum score of 100).

- Value (Stage 2), of moderate importance (e.g., an additional 30 points), concerns the detailed strategic, economic, and dissemination value of each artifact. For instance, we certainly seek relevant mature artifacts to add to the list, but we would rate low those logistically difficult or resource-prohibitive to implement. We recommend favoring artifacts that address clearly important aviation safety domains (air traffic, aircraft maintenance, design, manufacturing, human factors, etc.) and societal needs (commerce, the public, the military, etc.)

- Quality (Stage 3) involves a variety of internal and external quality features characterizing the artifact even further (e.g., design complexity, documentation, accuracy, and input data source availability). We recommend it carries the least score (the remaining 20 points).

To be effective, each stage should have a cut-off limit (denoted by X, Y and Z in Figure 1 and subject to the selection team's rigor) whereby artifacts are dropped if no longer meeting a certain minimum cumulative score.

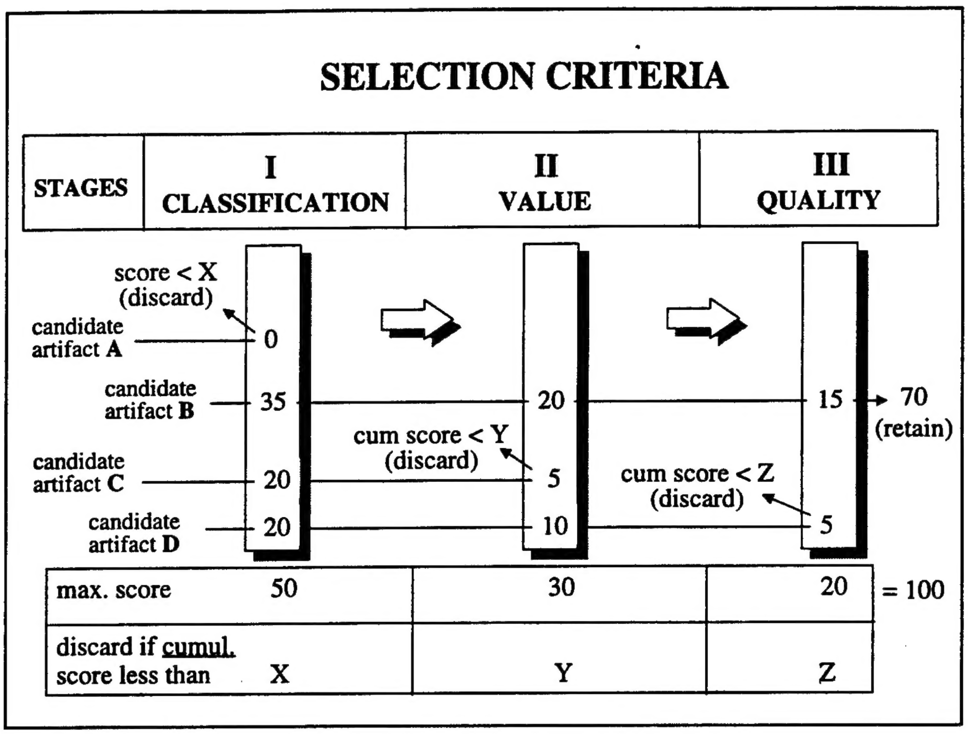

Figure 1. Selection criteria in three sequential stages.

2.2 Framework and Premises

Before conjecturing any criteria, we must use some sort of framework as a rational guide. That is, we must be equipped with enough constructs, assumptions, and boundaries about what we seek to launch. We prefer this approach because it instills logical discipline and rigor in the selection process from the start. Therefore, we propose "setting the stage" using the following premises:

- We define analysis to be a cognitive, logical, evaluative process synthesizing information into useful knowledge. Accordingly, we propose using this definition in assessing analytical artifacts. Thus, items declared non-analytical (whether aviation-related or not) should not be considered. Safety analysis means analysis that carries one or more of the following ulterior motives: (a) assessment, (b) prediction, (c) decision, or (d) design of some safety facet. That is, the intention of the artifact must go beyond just analyzing data accurately or effectively. Safety usefulness in these motives must be clearly implied in that intention. Only artifacts that fit this premise should receive high scores. Finally, aviation safety analysis means safety analysis dedicated clearly and deliberately to one or more facets of aviation.

- We assume we can find documented instances describing a wide range of analytical elements (e.g., methods, models, metrics, indicators, and tools). We choose to call these elements analytical artifacts or just artifacts. Thus, we plan to exclude entities such as books, reports, statements, and articles unless they describe an artifact. Excluded also are statements and reports using only prose to detail accidents and incidents. However, we would include analytical artifacts described in such statements.

- Before conducting our selection, we should be able to define each candidate artifact. For example, we should know readily whether it consists of an equation, metric, model, or software package by nature. We should also be able to tell whether it is safety-oriented or not. Finally, we should know its life cycle "maturity" status (research, prototype phase, deployment, etc.). This premise relies on the assumption that we cannot really evaluate something unless we define it first. Thus, artifacts missing one or more of the three basic definitions (analytical nature, safety relevancy, and maturity) do not fit our intended framework. We propose eliminating them from consideration. Therefore, we make classifying (defining) artifacts in Stage 1 mandatory, and we make value and quality (Stages 2 and 3, respectively) only ancillary to the selection process.

- Finally, we assume practitioners seeking artifacts proceed according to the usual sequence of (a) defining the safety problem, (b) sizing the required analysis, (c) identifying data sources, (d) adopting an analysis strategy, and (e) selecting and applying the appropriate analytical artifacts. Artifacts found non-compatible with this rational sequence are to be rated low.

Using this framework as a foundation, we can now mechanize the selection criteria process with sufficient rationality. Accordingly, the next three sections describe the process built into the three stages as shown in Figure 1. We provide a sample Individual Scoring Sheet in Appendix B and a Summary Scoring Sheet in Appendix C.

2.3 Classification

Because the framework emphasizes classification, we choose to label artifacts in three ways: safety relevancy level, analytical classes, and maturity status.

2.3.1 Safety relevancy levels

First, we propose three safety relevancy levels, defined as follows:

- Level III: the artifact is general, not specifically safety-oriented, but with potential use in safety analysis (e.g., a bayesian decision tree model that can easily be modified to satisfy a safety risk analysis).

- Level II: the artifact is safety-oriented but not specifically towards aviation. However, the item can be modified towards aviation safety (e.g., a cause and effect analysis tool used in nuclear reactor safety or a generic organizational behavior safety model that could be adapted for aviation).

- Level I: the artifact is explicitly oriented towards aviation safety (e.g., an aircraft collision risk model).

2.3.2 Analytical Classes

Next, we propose defining all types of analytical classes of artifacts including those that do not require safety to be their necessary inherent feature. They are

- Class E: generalized analytical methods, procedures, tools, or software with no intentional predisposition towards a specific domain like safety, quality, or reliability. Examples include generic statistical packages, commercial spreadsheets, tabulations, and mathematical formulas and equations, all of generalized meaning regardless of the field of application.

- Class D: individual measures, metrics, indicators, indices, and figures of merit with specific application intention. An example includes safety performance indicators. Class D artifacts most likely stand alone. However, they are often seen as part of a larger method, tool, or model. They are also often seen as part of larger collections of the same Class D.

- Class C: composite analytical methodologies and processes with specific application intention. A methodology consists of a rational process following a disciplined path where more than one analytical step is involved and where a particular objective is to be attained at the end of the process. Usually, methodologies involve Classes B and D artifacts. Examples include fault tree analysis, cause and effect diagrams, petri nets, and reliability charts directed at safety.

- Class B: analytical or simulation models with specific application intention. A model is an artifact enriched by a strong theoretical foundation, well frameworked within rational constructs (paradigms) and usually directed at a very specific purpose or application area. Examples include aircraft separation models, microburst weather simulation models, organizational safety models, and human factors simulation models.

- Class A: practical tools that are usually automated, documented artifacts with outputs directed at one or more aspect of a specific application. Tools are most likely operational (procedural) descendants of measures, models, and methodologies, binding one or more of them into a working version. Examples include a quality-control charting tool that tracks defects on a production line, a software package for fault tree analysis of aircraft systems, a petri net software package dedicated to timing faulty events in a network, and a software package to analyze operational errors and deviations data officially recorded in an air traffic control system.

2.3.3 Maturity Status

Finally, we propose defining the maturity status of the artifacts, as follows:

- State 4: the artifact has a reliable past record of accomplishment, evidence of validation and verification, and widespread use.

- State 3: the artifact is implemented but in limited or restricted use, has limited evidence of validation and verification, or has a limited past record of accomplishment.

- State 2: the artifact is beyond the research and development state, but in prototype form only at very few test sites or has limited evidence of validation and verification, with only broad plans for broadcasting and deployment.

- State 1: the artifact remains in research and development, in one or few incubators, with minimal validation and verification or with some broadcast in the literature with no specific plans for implementation and deployment.

2.3.3.1 Classification Tables

Once we accept these definitions as our starting point, we can combine them into a two-way criteria table (Table 1). Maximum scores are denoted as c_1, c_2, c_3, and so on, as shown in each cell. At this time, they remain numerically unassigned parameters, subject to the rigor of the selection team.

Maturity status weights are to be fixed as w_1, w_2, w_3, and w_4, as shown in Table 2. They also remain unassigned. To produce a final classification score, we recommend Table 1 scores be multiplied by these weights.

Table 1. Artifact Classification

<table>
<tr><th></th><th>Class A
Practical tools</th><th>Class B
Models</th><th>Class C
Methods</th><th>Class D
Individual measures, metrics, indicators</th><th>Class E
Generalized analytical artifacts</th></tr>
<tr><td>Level I
aviation safety oriented</td><td>highly desirable
(c_1)</td><td>desirable if not excessively complex
(c_3)
otherwise, moderately desirable
(c_4)</td><td>very desirable
(c_2)</td><td>very desirable if part of a collection of measures or included in a model or method
(c_2)
otherwise desirable
(c_3)</td><td rowspan="3">Undesirable by themselves
(c_7)
or
minimally desirable if readily modifiable for integration with other safety artifacts
(c_6)</td></tr>
<tr><td>Level II
safety oriented, but not necessarily towards aviation</td><td colspan="2">slightly desirable if substantial rework expected
(c_5)
otherwise moderately desirable
(c_4)</td><td colspan="2">desirable
(c_3)</td></tr>
<tr><td>Level III
general, not safety oriented</td><td colspan="4">undesirable by themselves
(c_7)
or
slightly desirable if readily modifiable for integration with other safety artifacts
(c_5)</td></tr>
</table>

Table 2. Maturity Status

maturity status	Multiplication factor to apply to Table 1 scores
State 4: in widespread use	w_1
State 3: limited, restricted use	w_2
State 2: working prototype only	w_3
State 1: research and development	w_4

2.4 Value

The value criteria address the strategic, economic, and informational advantages of the artifact. They are listed below with maximum scores shown in parentheses as v_1, v_2, v_3, and v_4. All scores remain numerically unassigned parameters, subject to the evaluation team's rigor.

a. Strategic advantage:

   1. The artifact applies clearly to one or more of the following aviation domains where safety is constantly an important issue: (a) aircraft operations, maintenance, design, and manufacturing; (b) air traffic; (c) aviation weather; (d) aviation human factors; and (e) aviation communications, navigation, and surveillance. Societal benefit ramifications are clearly implied in the scope and potential of the artifact. Examples of societal benefits include potential contributions towards private, national, and international commercial aviation, the military, and general public (v_1).
   2. Universality of application is present. The artifact carries a global (aviation-wide) theme. It can be seen as universal to many organizations and interest groups. It also carries few or no local/technical constraints that may prevent more widespread application (v_2).

b. Economic advantage:

   Implementation, deployment, and usage costs are low, learning curve short, and labor hours low relative to the potential benefits derived from applying the artifact (v_3).

c. Dissemination, training, and usability advantage:

   Artifact dissemination, training, and usage are straightforward. External consultancy is minimal. The artifact fits the analyst's customary sequence of (a) defining the problem, (b) sizing the analysis, (c) identifying data sources, (d) adopting an analysis strategy, and (e) selecting and applying the appropriate artifacts (v_4).

d. Disposition of ratings labeled unknown:

Because we believe that the detailed properties of the artifact may not be known first hand even though they may be present, we propose using a null score (0) annotated with a rating of unknown. The number of unknowns should be reported as shown in both the Individual Scoring Sheet (Appendix B) and the Summary Scoring Sheet (Appendix C).

2.5 Quality

By artifact quality, we mean several external and internal features. They are listed in the subparagraphs below with maximum scores shown in parentheses as q_1, q_2, q_3, and so on. All scores remain open based on the evaluation team's rigor.

a. External qualities:

1. The computation requirements are technically reasonable and feasible (q_1). Data are physically available (q_2), and data can easily be obtained to make them work (q_3).

2. Documentation is sufficiently informative. The artifact is described well in one or more sources, several expert points of contact exist in the community, and language and semantics are well known and recognizable in the safety community (q_4); and obscure notation and mathematical expressions are low (q_5).

3. Flexibility is present. The artifact is easily modifiable to fit into a larger context of safety analysis (q_6).

4. Independence is present. The artifact can stand alone, is results-wise, or has low or no analytical dependency on other successor or predecessor artifacts to be useful (q_7).

b. Internal qualities:

1. The scope is explicitly understood. Analytical objectives, assumptions, and limitations are coherent, without contradiction or deviation from each other (q_8).

2. The design complexity is low, or if high, documentation is properly modularized and understandable. In addition, results are clearly displayed and easily accessible for interpretation (q_9).

3. Accuracy is explicit because the artifact has provisions for reporting error tolerance levels in the results (q_{10}).

c. Disposition of ratings labeled unknown:

Because we believe that the detailed properties of the artifact may not be known first hand even though they may be present, we propose using a null score (0) annotated with a rating of unknown. The number of unknowns should be reported as shown in both the Individual Scoring Sheet (Appendix B) and the Summary Scoring Sheet (Appendix C).

Appendix A
Sample of Artifacts

(Source of Information: Office of System Safety, Federal Aviation Administration)

Accident/Incident Report (ADREP) System

The International Civil Aviation Organization (ICAO) gathers information on aircraft incidents considered important for safety and prevention. Member nations, typically developed nations such as the United States, Canada, Japan, the United Kingdom, and other European countries submit information to an ICAO compiler who enters the data using a pre-coded checklist form. Two forms are used: The Preliminary Report, which is used for accidents only and the Accident Data Report, which contains causes and safety recommendations. Each form contains a short answer section, sequence of events, and a narrative description. The ADREP system analyzes the accident or incident by a sequence of events, each detailed with up to five descriptive factors that identify accident or incident causes. Each descriptive factor is supported by up to three explanatory factors, which describe why it occurred. The ADREP system publishes bi-monthly summaries and annual statistics for broad categories of data.

Air Carrier Assessment Tool (ACAT)

A primary tool used in the Air Transportation Oversight System to develop a comprehensive surveillance plan (CSP) for an air carrier. The ACAT assesses the 88 air carrier system elements using a series of risk indicators. The CSP is developed annually and revised throughout the year to retarget surveillance based upon the continuous analysis of data and the identification of emerging safety trends. The CSP completely replaces current National Program Guidelines required for surveillance and planned surveillance programs.

Aircraft Movement Area Safety System (AMASS)

The AMASS integrates information from Airport Surface Display Equipment radar and terminal area surveillance radar to identify and alert controllers to runway incursions. The frequency of such alerts represents one safety performance measure. The system records much more information about the operation of aircraft on the airport surface that can be used to identify the situation. It does not trigger an alert but may still be indicative of an incipient safety problem.

Airspace Occupancy Model and Airspace Encounter Model (AOM and AEM)

The AOM and AEM are tools developed by the FAA-sponsored National Center of Excellence for Aviation Operations Research. AOM estimates three-dimensional airspace occupancy and provides inputs to the AEM, which models aircraft encounters, generating data on encounter geometries. Both models generate results mathematically, avoiding problems inherent in time-step simulation models. Airspace region of almost any shape and aircraft encounters of almost any type can be modeled.

Aviation Performance Measuring System (APMS)

The APMS is an R&D project managed by NASA to develop the next generation of tools for Flight Operational Quality Assurance. APMS will combine "special events" data with "atypical flight data" to isolate those events/phase of flight that were contrary to Standard Operating Procedures and atypical. The APMS includes a "suite of integrated tools" such as screening for special events, statistical analysis, and database exploration for atypical analysis and database exploration for atypical flights and flight simulation/animation.

Aviation System Indicators

The FAA-developed aviation system and environment indicators provide a comprehensive view of the National Aviation System operation and environment. To expedite the development of these indicators, an executive level Task Force with representatives from all major program areas was formed. The Task Force identified an initial set of system and environmental indicators, which have been modified over time to now include 25 system indicators and 12 environment indicators. New indicators will be added over time because of an ongoing review process to assess the status of aviation system performance. Current indicators will be modified and refined, as appropriate, to ensure their continuing adequacy and validity as measures of system performance. Actual monthly rates indicate the number of accidents or incidents that occurred during that month divided by appropriate measure of activity (e.g., flight hours or departures). A 12-month moving average is used.

Flight Track Analysis System (FTAS)

The FTAS utilizes the FAA Automated Radar Terminal System track data to provide a graphical display of the routes followed by each aircraft in a terminal area. It can replay a period of time at varying speeds and assign different colors to various classes of aircraft to facilitate visualization of the track patterns. In addition to its visualization capabilities, it can generate statistical charts and reports and provide the user with the ability to extract altitude and speed profiles or monitor the distribution of aircraft crossing defined locations. It is primarily used to generate the inputs into airspace and airport simulation models or aircraft noise analysis models, although it also provides a useful capability to explain airspace procedures to airport management and to the public.

Human Performance Models (HPMs)

HPMs are quantitative, analytic, or computer-based models that represent job-related behavior of the operators or maintainers of complex dynamic systems. There are two types: Output models that link input states to output states and do not address process and Process models that are theories of how people perform certain tasks.

Reduced Aircraft Separation Risk Assessment Model (RASRAM)

RASRAM links aircraft separation with quantitative safety risk. The model evaluates safety risks for a variety of flight scenarios relating to final approach, landing, and rollout for parallel and single runways. RASRAM computes the increase in risk of reduced separation operations during instrument meteorological conditions, considering procedural and technological changes. RASRAM was developed for three scenarios: 1) runway occupancy, 2) wake vortex encounter, and 3) blunder scenarios. The result for each scenario is a consolidated risk of incident and accident from all sources applicable to the scenario. Within a scenario, the RASRAM method incorporates fault trees and event trees, fixed probabilities, and time dependent probabilities.

Systematic Air Traffic Operation Research Initiative (SATORI)

The SATORI is an extensive tool developed by the Civil Aeromedical Institute to support accident and incident investigation. It utilizes the radar track data to recreate the radar display as seen by the controller at the time. The radio frequency recordings are digitized and synchronized with the radar display so that the investigators can "relive" the event. The system provides users with the ability to reconfigure the display to examine aspects of interest as well as to obtain objective measures of controller actions. The system also has useful applications for training and research, particularly into human factors issues involving controller tasks.

Appendix B
Sample Individual Scoring Sheet

Artifact name:	**Source of artifact:**	**Scoring date:**

Classification			
Safety level (I, II, or III)			
Analytical class (A, B, C, or D)			
Maturity status (1, 2, 3, or 4)			
score:		weight:	**total:** (score x weight) =

Value	score	unknown
Aviat. domains and society benefits		
Universality		
Economics		
Dissemination, training, usability		
total:		

Quality	score	unknown
Computation		
Documentation		
Flexibility		
Independence		
Scope		
Design		
Accuracy		
total:		

FINAL SCORE		
	score	unknowns
Sum of all 3 totals		

Appendix C
Sample Summary Scoring Sheet

	Rating								
	Classification			**Value**		**Quality**		**Total**	
Artifact name	label	raw score	weighted score	score	unknowns	score	unknowns	score	unknowns
XYZ	A-I-3								
ABC	B-III-2								
PQR	C-II-4								

June 17, 1999

DISTRIBUTION LIST – ACT-560

	Copies
ACL-1	Documentation Page Only
ACT-500	Original (J. Dunn)
ACT-71A	2 (Technical Center Library)
SVC122.10	4 (Headquarters Library)

1 copy each

AAR-1
AAR-2
ARR-100
ARR-200
ARR-400
AAT-1
ACT-1
ACT-2
ACT-500
ACT-510
ACT-520
ACT-530
ACT-540
ACT-550
ACT-560
AFS-1
AIT-1
AOZ-1
AOZ-2
APO-1
ARA-1
ASD-1
ASD-100
ASD-400
AST-1
AST-2
AST-100
AST-200
ASY-1
ASY-2
ASY-100
ASY-300
ATS-1
ATX-400
AVR-1